

Physiologically motivated multiplex Kuramoto model describes phase diagram of cortical activity

Maximilian Sadilek¹ and Stefan Thurner^{1,2,3,*}

¹*Section for Science of Complex Systems, Medical University of Vienna, Spitalgasse 23, A-1090, Vienna, Austria*

²*Santa Fe Institute, 1399 Hyde Park Road, New Mexico 87501, USA*

³*IIASA, Schlossplatz 1, A-2361 Laxenburg; Austria*

We derive a two-layer multiplex Kuramoto model from weakly coupled Wilson-Cowan oscillators on a cortical network with inhibitory synaptic time delays. Depending on the coupling strength and a phase shift parameter, related to cerebral blood flow and GABA concentration, respectively, we numerically identify three macroscopic phases: unsynchronized, synchronized, and chaotic dynamics. These correspond to physiological background-, epileptic seizure-, and resting-state cortical activity, respectively. We also observe frequency suppression at the transition from resting-state to seizure activity.

Fast electrochemical processes taking place on a complicated cytoarchitectural network structure render the human brain a highly complex dynamical system. Brain activity, as measured directly via EEG or MEG, or indirectly by means of MRI recordings, reveals characteristic macroscopic patterns such as oscillations in various frequency bands [1], synchronization [2–4], or chaotic dynamics [5]. Generally it is believed that macroscopic activity (involving $10^8 - 10^{11}$ neurons) is closely related to high-level functions such as cognition, attention, memory or task execution. To understand the mechanisms of this correspondence, both the overall network structure of the brain and the local properties of neural populations have to be taken into account [4, 6]. Regarding the latter, neural inhibition seems to be essential for cortical processing [7].

Two physiological phenomena received much attention lately in terms of a mathematical understanding: resting-state activity, and gamma oscillations. Resting-state activity is spontaneous, highly structured activity of the brain during rest, and can be described in terms of networks of simultaneously active brain regions [8, 9]. Models of resting-state networks often rely on anatomical networks derived from histological or imaging data, and on local interactions between populations of excitatory and inhibitory neurons [10–13]. Oscillatory neural activity in the gamma range (30 – 100 Hz) is potentially related to consciousness and the binding problem although its precise function remains unclear [14]. To understand the origin of gamma oscillations, two mechanisms have been proposed [15]. One describes interactions between inhibitory neurons together with an external driving force [16, 17]. The other mechanism is based on excitatory-inhibitory coupling with synaptic time delays [18–20]. The relation of gamma oscillations and inhibition is experimentally well established. In mice [21], rats [22] and humans [3, 23], a decrease of GABA-concentrations (Gamma-aminobutyric acid is the main inhibitory neurotransmitter in mammals) is accompanied by a strong attenuation of the gamma frequency

band and sometimes by epileptiform activity.

Many existing network models for resting-state activity and gamma oscillations are based on single-neuron local dynamics [10, 11, 16–19]. Since experimentally observed resting-state networks comprise individual regions containing about 10^8 to 10^9 individual neurons, we believe that a local description in terms of Wilson-Cowan equations is an attractive alternative. In this work we derive a simple two-layer multiplex model from classical physiological equations that is able to capture the main features of cortical activity such as oscillations, synchronization and chaotic dynamics. This model unifies the roles of neural network topology, synaptic time delays, and excitation/inhibition. It provides a closed framework for simultaneously understanding the origin of resting-state activity and gamma oscillations.

We consider N cortical regions indexed by $i = 1, \dots, N$, see Fig. 1(a). Each region is populated by ensembles of excitatory and inhibitory neurons (e.g. pyramidal cells and interneurons). We define the activity level of a region i as the fraction of firing excitatory (inhibitory) neurons of the total number of excitatory (inhibitory) neurons in that region at a unit time interval, and denote it by x_i (y_i). Neglecting for the moment interactions between different regions, we assume that individual cortical regions obey the Wilson-Cowan-type dynamics [20]

$$\begin{aligned} \dot{x}_i &= -x_i + S\left(a_i x_i - b_i y_i + \rho_i^{(x)}\right) \\ \dot{y}_i &= -y_i + S\left(c_i x_i - d_i y_i + \rho_i^{(y)}\right) \end{aligned} \quad (1)$$

where S is a sigmoidal response function, a_i and d_i are real-valued feedback parameters, and b_i and c_i are positive synaptic coefficients ($i = 1 \dots N$). $\rho_i^{(x)}$ and $\rho_i^{(y)}$ account for external inputs, e.g. from sensory organs. We now introduce interactions among regions of the network

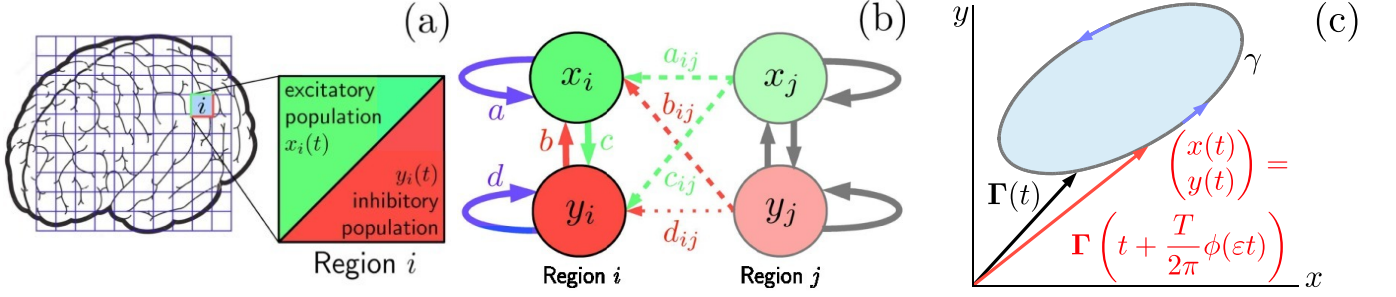


FIG. 1. (a) The cortical surface is divided into N macroscopic regions. Every region i (blue) comprises excitatory (green) and inhibitory (red) populations with activity levels $x_i(t)$ and $y_i(t)$, respectively. Activity levels quantify the ratio of firing neurons in the region at time t . (b) Region i (left) receives excitatory (green arrows) and inhibitory (red arrows) inputs from adjacent regions j (right) plus self-feedback (blue arrows). (c) Variable transformation from activity variables $x_i(t)$ and $y_i(t)$ to phase deviation variables $\phi_i(\varepsilon t)$. On a limit cycle γ , ε -perturbations of the (x, y) -dynamics at time t induce the phase deviations $\phi(\varepsilon t)$.

by replacing $\rho_i^{(\cdot)}$ in Eq. (1) by

$$\begin{aligned} \rho_i^{(x)} &\rightarrow \rho_i^{(x)} + \sum_{j \neq i} [a_{ij}x_j(t) - b_{ij}y_j(t - \tau)] \\ \rho_i^{(y)} &\rightarrow \rho_i^{(y)} + \sum_{j \neq i} [c_{ij}x_j(t - \tau) - d_{ij}y_j(t)] \end{aligned} \quad (2)$$

Here $a_{ij}, b_{ij}, c_{ij}, d_{ij}$ are positive synaptic coefficients linking regions i and j . τ accounts for transmission delays at inhibitory synapses (not to be confused with axonal conduction delays). In physiology, τ can be altered by changing the synaptic concentration of GABA [18, 19, 21, 22]. In the present model, we assume that τ is proportional to the average synaptic GABA concentration in the brain. To derive a multiplex Kuramoto model (MKM) from Eqs. (1) and (2), we make the following three assumptions:

(i) *Homogeneity*. Cortical regions exhibit nearly identical dynamical behavior. We therefore assume the following parameters to be constant across regions,

$$(a_i, b_i, c_i, d_i, \rho_i^{(x)}, \rho_i^{(y)}) = (a, b, c, d, \rho^{(x)}, \rho^{(y)}) + \mathcal{O}(\varepsilon) \quad (3)$$

for all i , up to small perturbations, denoted by ε .

(ii) *Stable local oscillations*. We choose the parameters $(a, b, c, d, \rho^{(x)}, \rho^{(y)})$ such that each uncoupled system Eq. (1), under the assumption given in Eq. (3), has a unique exponentially stable limit cycle $\gamma \in \mathbb{R}^2$. As a consequence, after a transient time solutions of Eq. (1) can be written as

$$\begin{pmatrix} x_i(t) \\ y_i(t) \end{pmatrix} = \Gamma(t + \varphi_i) \quad (4)$$

where $\Gamma(t)$ is an arbitrary solution of Eq. (1) on γ . φ_i accounts for specific initial values [24, 25]. Let T denote the period of Γ . We assume that the frequency $1/T$ lies in the physiological gamma range.

(iii) *Weak coupling*. Interactions between adjacent regions are weak, and inhibitory-inhibitory interactions are

very weak in the sense that,

$$(a_{ij}, b_{ij}, c_{ij}, d_{ij}) = \varepsilon(A_{ij}, B_{ij}, C_{ij}, 0) + \mathcal{O}(\varepsilon^2) \quad (5)$$

for all $i \neq j$. These assumptions are justified because the number of synaptic connections within a cortical region is much larger than between regions, and excitatory neurons outnumber inhibitory neurons by approximately one order of magnitude [26, 27].

Figure 1(b) summarizes the connectivity structure between regions i and j . Region i receives excitatory (green arrows) and inhibitory (red arrows) inputs plus feedback (blue arrows), magnitudes are indicated by the arrow labels. For the sake of clarity, arrows representing inputs of magnitude $\mathcal{O}(1)$, $\mathcal{O}(\varepsilon)$ and $\mathcal{O}(\varepsilon^2)$ are drawn in continuous, dashed and dotted style, respectively.

Under these assumptions, the system Eq. (1) with Eq. (2) is equivalent [28] to a two-layer MKM

$$\begin{aligned} \frac{d\phi_i}{dt} &= \omega_i + \frac{K}{\langle k \rangle} \sum_{j=1}^N A_{ij} \sin(\phi_j - \phi_i) \\ &+ \frac{K}{\langle k^{(\delta)} \rangle} \sum_{j=1}^N A_{ij}^{(\delta)} \sin(\phi_j - \phi_i - \delta) \end{aligned} \quad (6)$$

Here $\phi_i(t) \in \mathcal{S}^1$ describes the deviations from the uncoupled phases $\theta_i(t) \equiv 2\pi/T(t + \varphi_i)$ that are associated with solutions of the uncoupled system Eq. (4). Accordingly, $d\phi_i/dt$ describes the deviations from the uncoupled oscillation frequency $2\pi/T$. Time t has been rescaled [28]. A_{ij} is the adjacency matrix of the excitatory-excitatory interaction network as defined in Eq. (5), $A_{ij}^{(\delta)} \equiv 1/2(B_{ij} + C_{ij})$ accounts for the interaction between excitatory and inhibitory populations, and $\langle k \rangle \equiv 1/N \sum_{i,j} A_{ij}$, and $\langle k^{(\delta)} \rangle \equiv 1/N \sum_{i,j} A_{ij}^{(\delta)}$, are the corresponding average degrees.

δ is a phase shift parameter related to the time delay τ via $\delta \equiv (2\pi/T)\tau$. K is a global coupling constant that we assume to be proportional to the cerebral blood

flow. This is reasonable because the latter is strongly correlated with the connection strengths of functional networks reconstructed in magnetic resonance imaging [29]. ω_i , the so-called *natural frequencies* of the MKM, are the constant contribution to the frequency deviations $d\phi_i/dt$. We take ω_i from a symmetric, unimodal random distribution $g(\omega)$, with mean ω_0 . Since the 1-parameter family of rotating-frame transformations $\phi_i(t) \rightarrow \phi_i(t) - \Omega t$, $\omega_i \rightarrow \omega_i - \Omega$, leave Eq. (6) invariant for any Ω , without loss of generality we assume, $\omega_0 = 0$ and $\delta \in [0, \pi]$. Note that for each solution $\phi_i(t)^*$ with $\delta^* \in [\pi, 2\pi]$, there exists a solution $\phi_i(t)$ with $\delta = 2\pi - \delta^* \in [0, \pi]$ and $\phi_i(t) = -\phi_i(t)^*$. Physiological processes changing δ , K , and ω_i occur on a much slower timescale than neural activity.

It is known that weakly coupled, nearly identical limit-cycle oscillators can be described in terms of phase variables [25, 30–32]. However, in terms of the new variables $(\phi_1 \dots \phi_N)$, interactions between cortical regions take place on two independent layers representing excitatory-excitatory and excitatory-inhibitory coupling, respectively, and the complicated connectivity structure of Fig. 1(b) reduces to a simple two-layer multiplex structure. Figure 1(c) shows the variable transformation from activity variables x and y , to the phase variable ϕ for any cortical region. In the unperturbed case, $\varepsilon = 0$, the limit cycle γ is parametrized by $\Gamma(t) = (x(t), y(t))$. Since γ is exponentially stable, ε -perturbations of activity dynamics $t \mapsto (x(t), y(t))$ lead to phase deviations $\Gamma(t + T/(2\pi)\phi(\varepsilon t))$. For $\delta = 0$ we recover the Kuramoto model on a single network, see [28, 33–38].

Synchronization. We define the order parameter [31–33, 39]

$$r(t) \equiv \left| \frac{1}{N} \sum_{j=1}^N \exp[i\phi_j(t)] \right|. \quad (7)$$

It takes values between 0 (no synchronization) and 1 (full synchronization) [25]. Let \bar{r} denote its time average $\bar{r} \equiv \langle r(t) \rangle_t$.

Chaotic dynamics. The instantaneous largest Lyapunov exponent is given by

$$\lambda_{max}(t) \equiv \lim_{d_0 \rightarrow 0} \frac{1}{t - t_0} \ln \frac{d(t)}{d_0}, \quad (8)$$

where $d(t) = \|\phi(t) - \phi^p(t)\|$ measures the separation between a reference trajectory $\phi(t)$ and a perturbed one $\phi^p(t)$. d_0 is the initial separation at t_0 , and $\|\cdot\|$ is the 1-norm [28]. For large times, $\lambda_{max}(t)$ approaches the “true” largest Lyapunov exponent, $\bar{\lambda}_{max}$.

Average frequency deviation. We look at average frequency deviations across all regions,

$$\Omega \equiv \frac{1}{N} \sum_{i=1}^N \frac{d\phi_i}{dt}, \quad (9)$$

once a stationary state is reached.

Equation (6) is integrated with a standard 4th-order Runge-Kutta algorithm with 500 time steps of size $dt = 0.1$. The system size is $N = 100$, both layers are chosen to be Erdős-Rényi networks with $p = 0.06$. Natural frequencies ω_i are taken from a standard normal distribution, initial phase deviations $\phi_i(0)$ from the interval $[0, 2\pi]$. The first 250 time steps are discarded to exclude transient effects. For the remaining time steps, \bar{r} , $\bar{\lambda}_{max}$, and Ω are evaluated. All results are averaged over 100 identical, independent runs with different realizations of the initial conditions.

Synchronization. We find that synchronization \bar{r} depends on the coupling strength K , and phase shift δ , Fig. 2(a). For $\delta = 0$, we expect [28] a transition from an unsynchronized to a synchronized state at a critical value $K_c \approx 0.74$, which is confirmed by our simulations, Fig. 2(a). With $\delta > 0$, stronger coupling K is required for this transition to occur. Above a value of approximately $\delta = \pi/2$, a global synchronized state ceases to exist.

Chaotic dynamics. Above the synchronization threshold, $K > K_c$, synchronization and chaotic dynamics are mutually exclusive, see Fig. 2(b). For small values of δ , there exists a small chaotic region ($\bar{\lambda}_{max} \approx 0.1$) at the Kuramoto transition, in agreement with the well-known results for $\delta = 0$, [28]. This region is expanding with increasing values of δ . At the boundary to the synchronized region, increasingly large values of $\bar{\lambda}_{max}$ are obtained. $\bar{\lambda}_{max}$ peaks at $\bar{\lambda}_{max} \approx 1$, for $\delta \approx \pi/2$. In the unsynchronized region, $K < K_c$, the dynamics is not chaotic, $\bar{\lambda}_{max} = 0$. Figure 2(c) integrates both results (synchronization and chaotic dynamics) into a schematic phase diagram that clearly exhibits three phases.

Spectral properties. Figure 3 (a)-(c) shows the stationary distributions of frequency deviations $d\phi_i/dt$ for selected values in the (K, δ) -plane. For $K = 0$, the distributions are practically identical for different values of δ , Fig. 3(a). For $K = 2.5$, at $\delta = 0$, a synchronization peak appears close to frequency zero. With increasing δ , this peak moves towards increasingly negative values, until $\delta \approx 3\pi/8$. Between $\delta \approx 3\pi/8$ and $3\pi/4$, the distribution is rapidly becoming broader and shifts towards positive values. After reaching a maximum at $\delta \approx 3\pi/4$, it is finally centered around zero again, Fig. 3(b). $K = 5$, is similar, however larger positive and negative values for $d\phi_i/dt$ occur, Fig. 3(c).

Figure 2(d) shows the average frequency deviation Ω as a function of K and δ . As expected [28], we find frequency suppression associated with synchronization in the region of large K and small δ , but also for large K and intermediate δ . For fixed K , maximal frequency suppression occurs at $\delta \approx 3\pi/8$. For large K and large δ (chaotic regime) we find slightly positive Ω .

In several variants of single-layer Kuramoto models with a phase shift or a time delay, frequency suppression

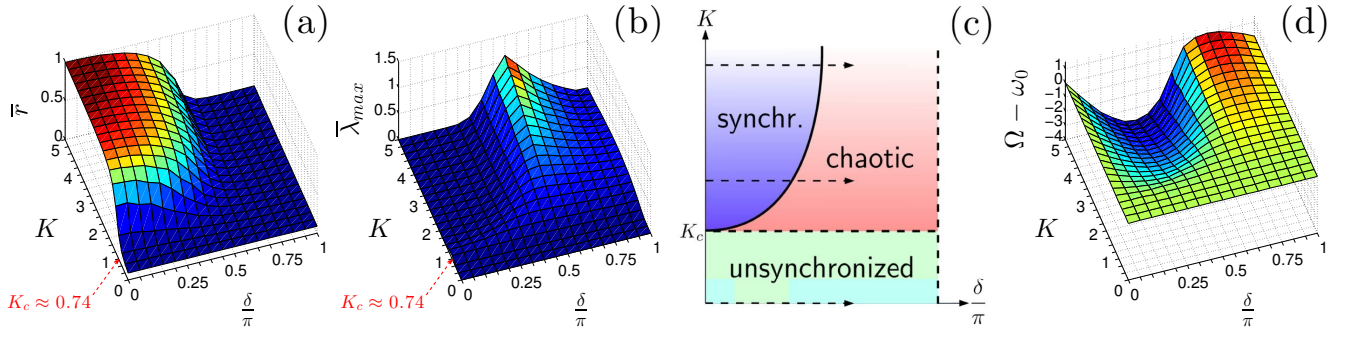


FIG. 2. (a) Order parameter \bar{r} and (b) largest Lyapunov exponent $\bar{\lambda}_{max}$ identify the regions of synchronization and of chaotic dynamics in the (K, δ) -plane. (c) Schematic phase diagram. Dashed arrows indicate directions along which distributions of frequency deviations were evaluated in Figs. 3(a)-(c). (d) Average frequency deviations Ω in the (K, δ) -plane.

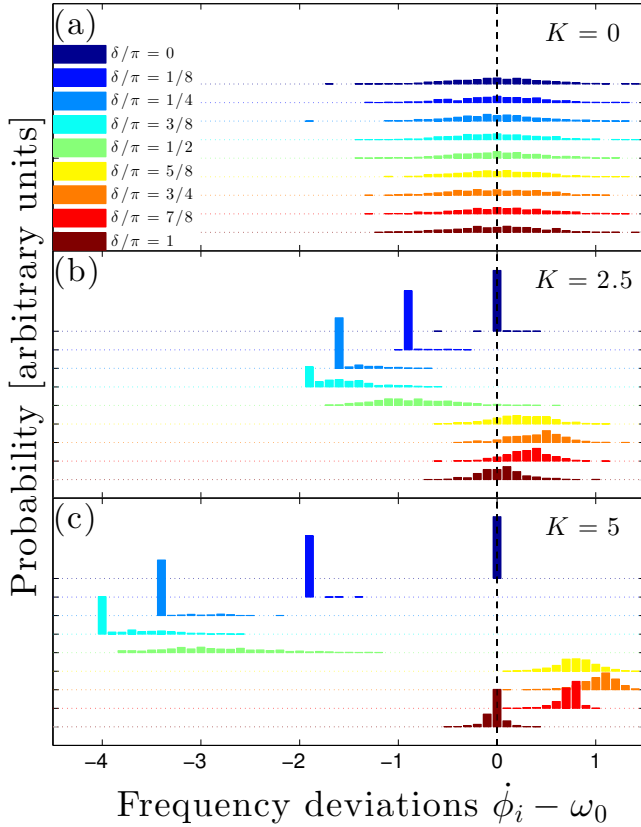


FIG. 3. Stationary distribution of frequency deviations $d\phi_i/dt$ for different δ values (colors) for (a) subcritical, (b) critical and (c) supercritical values of K , respectively. $K_c \approx 0.74$.

sion appears [36, 38, 40, 41]. In addition, [38] mentions chaotic behavior. However, the existence of the phase diagram with the three distinct macroscopic phases can not be inferred from any of those models to the best of our knowledge.

We tested the model for robustness with respect to the particular choice of parameters. As suggested by various brain atlases and cortical parcellation schemes, a number of $N \approx 100 - 150$ cortical regions seems rea-

sonable [42, 43]. We tested up to $N = 500$ and found no deviations from the presented qualitative picture. For the link density p , we find that as long as it exceeds the percolation threshold, $p > \log(N)/N$, differences in simulations are marginal. Finally, we observe that like in the original Kuramoto model [31, 32], for different natural frequency distributions $g(\omega)$ the qualitative behaviour remains practically unchanged as long as $g(\omega)$ is unimodal and symmetric.

Summarizing, we can show mathematically that a set of weakly coupled Wilson-Cowan oscillators on a cortical network with a synaptic time delay between excitatory and inhibitory neural populations is identical to a simple Kuramoto-type phase model on a two-layer multiplex network. Numerical investigations of this model reveal the presence of three distinct macroscopic phases in the space of control parameters K (associated with cerebral blood flow) and δ (associated with synaptic GABA concentration). For couplings $K < K_c$, activities of individual cortical regions show independent oscillatory behavior (unsynchronized). Frequencies are distributed symmetrically around an average frequency that we assume to be located in the physiological gamma range. This dynamical state corresponds to “background activity” of the brain. For $K > K_c$, two phases are possible: for small δ , the system becomes synchronized, which corresponds to “epileptic seizure activity” in physiology. For large δ , synchronized activity only appears in metastable clusters; the system is chaotic in general. We identify this phase with “resting-state activity” in the brain. An important property of the present model is that the average oscillation frequency is shifted towards lower values when crossing the boundary to the synchronized phase. This could explain the experimental fact [2, 21–23] that a decrease of the GABA concentration in the resting-state both triggers the appearance of epileptiform slow waves and diminishes gamma activity in the brain.

We acknowledge financial support from EC FP7 projects LASAGNE, agreement no. 318132, and MULTIPLEX, agreement no. 317532.

* stefan.thurner@meduniwien.ac.at

- [1] G. Buzsáki and A. Draguhn, *Science* **304**, 1926 (2004).
- [2] M. Eisenstein, *Nature* **511**, 4 (2014).
- [3] D. A. Lewis, T. Hashimoto, and D. W. Volk, *Nat. Rev. Neurosci.* **6**, 312 (2005).
- [4] F. Varela, J. P. Lachaux, E. Rodriguez, and J. Martinerie, *Nat. Rev. Neurosci.* **2**, 229 (2001).
- [5] C. J. Stam, *J. Clin. Neurophysiol.* **116**, 2266 (2005).
- [6] E. Bullmore and O. Sporns, *Nat. Rev. Neurosci.* **10**, 186 (2009).
- [7] J. S. Isaacson and M. Scanziani, *Neuron* **72**, 231 (2011).
- [8] B. B. Biswal, *Neuroimage* **62**, 938 (2012).
- [9] G. Deco, V. K. Jirsa, and A. R. McIntosh, *Nat. Rev. Neurosci.* **12**, 43 (2010).
- [10] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns, *Proc. Nat. Acad. Sci. USA* **104**, 10240 (2007).
- [11] A. Ghosh, Y. Rho, A. R. McIntosh, R. Kötter, and V. K. Jirsa, *PLoS Comput. Biol.* **4**, e1000196 (2008).
- [12] G. Deco, V. Jirsa, A. R. McIntosh, O. Sporns, and R. Kötter, *Proc. Nat. Acad. Sci. USA* **106**, 10302 (2009).
- [13] J. Cabral, E. Hugues, O. Sporns, and G. Deco, *Neuroimage* **57**, 130 (2011).
- [14] W. Singer and C. M. Gray, *Annu. Rev. Neurosci.* **18**, 555 (1995).
- [15] G. Buzsáki and X. J. Wang, *Annu. Rev. Neurosci.* **35**, 203 (2012).
- [16] M. A. Whittington, R. D. Traub, and J. G. R. Jefferys, *Nature* **373**, 612 (1995).
- [17] X. J. Wang and G. Buzsáki, *J. Neurosci.* **16**, 6402 (1996).
- [18] N. Brunel and X. J. Wang, *J. Neurophysiol.* **90**, 415 (2003).
- [19] C. Geisler, N. Brunel, and X. J. Wang, *J. Neurophysiol.* **94**, 4344 (2005).
- [20] H. R. Wilson and J. D. Cowan, *Biophys. J.* **12**, 1 (1972).
- [21] E. O. Mann and I. Mody, *Nat. Neurosci.* **13**, 205 (2010).
- [22] A. V. Medvedev, *Brain Res. Bull.* **58**, 115 (2002).
- [23] S. D. Muthukumaraswamy, R. A. E. Edden, D. K. Jones, J. B. Swettenham, and K. D. Singh, *Proc. Nat. Acad. Sci. USA* **106**, 8356 (2009).
- [24] R. M. Borisyuk and A. B. Kirillov, *Biol. Cybern.* **66**, 319 (1992).
- [25] F. C. Hoppensteadt and E. M. Izhikevich, *Weakly connected neural networks* (Springer Series Applied Mathematical Sciences, New York, 1997), Vol. 126.
- [26] A. Fairen, J. DeFelipe, and J. Regidor, *Cereb. Cortex* **7**, 347 (1984).
- [27] J. DeFelipe, I. Fariñas, *Prog. Neurobiol.* **39**, 563 (1992).
- [28] See supplemental material for details concerning the derivation of the multiplex Kuramoto model, the case $\delta = 0$ and the calculation of the instantaneous largest Lyapunov exponent.
- [29] T. Tsurugizawa, L. Ciobanu, and D. Le Bihan, *Proc. Nat. Acad. Sci. USA* **110**, 11636 (2013).
- [30] A. Winfree, *J. Theor. Biol.* **16**, 15 (1967).
- [31] Y. Kuramoto, *Self-entrainment of a population of coupled non-linear oscillators*, in *Int. Symp. on Math. Prob. in Theor. Phys.*, ed. H. Araki (Springer Series Lecture Notes in Physics, Berlin Heidelberg, 1975), Vol. 39.
- [32] Y. Kuramoto, *Prog. Theor. Phys. Supp.* **79**, 223 (1984).
- [33] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, *Phys. Rep.* **469**, 93 (2008).
- [34] A. C. Kalloniatis, *Phys. Rev. E* **82**, 066202 (2010).
- [35] G. Miritello, A. Pluchino, and A. Rapisarda, *Physica A* **388**, 4818 (2009).
- [36] E. Niebur, H. G. Schuster, and D. M. Kammen, *Phys. Rev. Lett.* **67**, 2753 (1991).
- [37] M. K. S. Yeung and S. H. Strogatz, *Phys. Rev. Lett.* **82**, 648 (1999).
- [38] V. Nicosia, M. Valencia, M. Chavez, A. Díaz-Guilera, and V. Latora, *Phys. Rev. Lett.* **110**, 174102 (2013).
- [39] J. A. Acebrón, L. L. Bonilla, C. J. P. Vicente, F. Ritort, and R. Spigler, *Rev. Mod. Phys.* **77**, 137 (2005).
- [40] V. H. P. Louzada, N. A. M. Araújo, J. S. Andrade, Jr., and H. J. Herrmann, *Nat. Sci. Rep.* **2** (2012).
- [41] M. Y. Choi, H. J. Kim, D. Kim, and H. Hong, *Phys. Rev. E* **61**, 371 (2000).
- [42] K. Zilles and K. Amunts, *Nat. Rev. Neurosci.* **11**, 139 (2010).
- [43] D. C. Van Essen, M. F. Glasser, D. L. Dierker, J. Harwell, and T. Coalson, *Cereb. Cortex* **22**, 2241 (2011).
- [44] S. H. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering* (Perseus Books Group, New York, 1994).

Supplemental Material

DERIVATION OF THE MULTIPLEX KURAMOTO MODEL

In this section, we derive the MKM, Eq. (6), from nearly identical, weakly coupled Wilson-Cowan oscillators, see Eqs. (1) - (5). The main idea is the following: Starting from uncoupled oscillators having exponentially stable orbits, the introduction of weak coupling between them only affects their phases (and not their frequencies or amplitudes). This allows for a transformation from activity- to phase deviation variables. The special structure of the Wilson-Cowan equations then accounts for the multilayer structure of the MKM.

For compact notation, we define $\mathbf{X}_i \equiv (x_i, y_i)^\top$ and $\mathbf{X} \equiv (\mathbf{X}_1, \dots, \mathbf{X}_N)$, where $^\top$ denotes transposition. Assumptions expressed in Eqs. (3) and (5), allow us to write Eq. (1) with the interactions introduced by Eq. (2) in the form

$$\frac{d\mathbf{X}_i(t)}{dt} = \mathbf{F}[\mathbf{X}_i(t)] + \varepsilon \delta \mathbf{F}_i[\mathbf{X}_i(t)] + \varepsilon \mathbf{G}_i[\mathbf{X}(t), \mathbf{X}(t - \tau)] \quad , \quad (\text{S10})$$

up to terms of order $\mathcal{O}(\varepsilon^2)$. Here \mathbf{F} constitutes the uncoupled part of the dynamics, and $\delta \mathbf{F}_i$ accounts for small perturbations in the parameters $a, b, c, d, \rho^{(x)}$, and $\rho^{(y)}$. \mathbf{G}_i encodes the contributions from weak coupling. Terms of order $\mathcal{O}(\varepsilon^2)$ will be omitted in all subsequent expressions.

Assumption from Eq. (4) states that solutions $\mathbf{X}_i^*(t)$ of the uncoupled system

$$\frac{d\mathbf{X}_i(t)}{dt} = \mathbf{F}[\mathbf{X}_i(t)]$$

approach a unique, exponentially stable limit cycle $\gamma \subset \mathbb{R}^2$. This means that after a transient phase we can write $\mathbf{X}_i^*(t) = \mathbf{\Gamma}(t + \varphi_i)$. We define the phase of such a solution by the mapping $\theta: \mathbf{X}_i^*(t) \mapsto 2\pi/T(t + \varphi_i) \in \mathcal{S}^1$, uniquely assigning a point on the unit cycle to each point of the solution, where T is the period of $\mathbf{\Gamma}$.

Since solutions $\mathbf{X}_i^*(t)$ of the uncoupled system are exponentially orbitally stable, we can write solutions of the weakly coupled system Eq. (S10) as

$$\mathbf{X}_i(t) = \mathbf{\Gamma} \left[t + \frac{T}{2\pi} \phi_i(\varepsilon t) \right] + \varepsilon \mathbf{P}_i \left[t + \frac{T}{2\pi} \phi_i(\varepsilon t) \right] \quad , \quad (\text{S11})$$

where $\phi_i(t) \in \mathcal{S}^1$ is the deviation from the phase $\theta_i(t) \equiv \theta(\mathbf{X}_i^*(t)) = 2\pi/T(t + \varphi_i)$, see Fig. 1(c). $\varepsilon \mathbf{P}_i$ accounts for the effects of the ε -perturbation of the invariant manifold $\gamma^N \subset \mathbb{R}^{2N}$, where γ^N denotes the N -th Cartesian power of γ , [25]. Initial values of ϕ_i are given by

$$\phi_i(0) = \frac{2\pi}{T} \varphi_i \quad .$$

Differentiating Eq. (S11) with respect to t and using Eq. (S10), we arrive at the linear inhomogeneous equation

$$\frac{d\mathbf{v}_i(t, \phi_i)}{dt} = \mathbf{M}(t, \phi_i) \mathbf{v}_i(t, \phi_i) + \mathbf{m}_i(t, \phi) \quad , \quad (\text{S12})$$

where $\phi \equiv (\phi_1(\varepsilon t), \dots, \phi_N(\varepsilon t))$ and

$$\begin{aligned} \mathbf{v}_i(t, \phi_i) &\equiv \mathbf{P}_i \left(t + \frac{T}{2\pi} \phi_i \right) \quad , \\ \mathbf{M}(t, \phi_i) &\equiv D\mathbf{F} \left(t + \frac{T}{2\pi} \phi_i \right) \quad , \\ \mathbf{m}_i(t, \phi) &\equiv \mathbf{G}_i \left(t + \frac{T}{2\pi} \phi, t - \tau + \frac{T}{2\pi} \phi \right) + \delta \mathbf{F}_i \left(t + \frac{T}{2\pi} \phi_i \right) - \mathbf{F} \left(t + \frac{T}{2\pi} \phi_i \right) \frac{d\phi_i}{d(\varepsilon t)} \quad . \end{aligned}$$

Here $D\mathbf{F}$ denotes the Jacobian matrix of \mathbf{F} , and we have used the abbreviations $\mathbf{F}(t) \equiv \mathbf{F}(\mathbf{\Gamma}(t))$, $\delta \mathbf{F}_i(t) \equiv \delta \mathbf{F}_i(\mathbf{\Gamma}(t))$, and $\mathbf{G}_i(t, t - \tau) \equiv \mathbf{G}_i(\mathbf{\Gamma}(t), \mathbf{\Gamma}(t - \tau))$ for clarity of notation. Now, since both solutions $\mathbf{v}_i(t, \phi_i)$ of Eq. (S12) and $\mathbf{w}(t, \phi_i)$ of the adjoint *homogeneous* system

$$\frac{d\mathbf{w}(t, \phi_i)}{dt} = -\mathbf{M}(t, \phi_i)^\top \mathbf{w}(t, \phi_i) \quad (\text{S13})$$

are periodic, nontrivial and unique by construction, the orthogonality condition

$$\frac{1}{T} \int_0^T \mathbf{w}(t, \phi_i)^\top \mathbf{m}_i(t, \phi) dt = 0 \quad , \quad (\text{S14})$$

follows (Fredholm alternative). If we normalize the solution $\mathbf{w}(t, \phi_i)$ of Eq. (S13) by the condition

$$\frac{1}{T} \int_0^T \mathbf{w}(t, \phi_i)^\top \frac{T}{2\pi} \mathbf{F} \left(t + \frac{T}{2\pi} \phi_i \right) dt = 1 \quad , \quad (\text{S15})$$

insert the definition of $\mathbf{m}_i(t, \phi)$ into Eq. (S14), and use Eq. (S15) we obtain

$$\frac{d\phi_i}{d(\varepsilon t)} = \frac{1}{T} \int_0^T \mathbf{w}(t, \phi_i)^\top \left[\delta \mathbf{F}_i \left(t + \frac{T}{2\pi} \phi_i \right) + \mathbf{G}_i \left(t + \frac{T}{2\pi} \phi, t - \tau + \frac{T}{2\pi} \phi \right) \right] dt \quad .$$

Due to Eq. (S13) and the special form of $\mathbf{M}(t, \phi_i)$ we have $\mathbf{w}(t, \phi_i) = \mathbf{w} \left(t + \frac{T}{2\pi} \phi_i, 0 \right)$. If we define $\mathbf{Q}(t) \equiv \mathbf{w}(t, 0)$ and make the substitutions $t \rightarrow t + \frac{T}{2\pi} \phi_i$ and $\varepsilon t \rightarrow t$, we obtain

$$\frac{d\phi_i}{dt} = \frac{1}{T} \int_0^T \mathbf{Q}(t)^\top \left\{ \delta \mathbf{F}_i \left[\mathbf{\Gamma} \left(t \right) \right] + \mathbf{G}_i \left[\mathbf{\Gamma} \left(t + \frac{T}{2\pi} (\phi - \phi_i) \right), \mathbf{\Gamma} \left(t - \tau + \frac{T}{2\pi} (\phi - \phi_i) \right) \right] \right\} dt \quad , \quad (\text{S16})$$

where we expanded the shortened notation for \mathbf{F} , $\delta \mathbf{F}_i$ and \mathbf{G}_i again. Making use of the structure of \mathbf{G}_i , see Eqs. (1), (2), (S10), we can rewrite Eq. (S16) as

$$\frac{d\phi_i}{dt} = \omega_i + \sum_{j \neq i}^N A_{ij} H_{EE}(\phi_j - \phi_i) + B_{ij} H_{EI}(\phi_j - \phi_i - \delta) + C_{ij} H_{IE}(\phi_j - \phi_i - \delta) \quad , \quad (\text{S17})$$

where δ is a phase shift parameter related to the time delay τ through $\delta = (2\pi/T)\tau$, and

$$\begin{aligned} \omega_i &\equiv \frac{1}{T} \int_0^T \mathbf{Q}(t)^\top \delta \mathbf{F}_i [\mathbf{\Gamma}(t)] dt \\ H_{EE}(\chi) &\equiv \frac{1}{T} \int_0^T Q_1(t) S' \left(ax(t) - by(t) + \rho^{(x)} \right) x \left(t + \frac{T}{2\pi} \chi \right) dt \\ H_{EI}(\chi) &\equiv -\frac{1}{T} \int_0^T Q_1(t) S' \left(ax(t) - by(t) + \rho^{(x)} \right) y \left(t + \frac{T}{2\pi} \chi \right) dt \\ H_{IE}(\chi) &\equiv \frac{1}{T} \int_0^T Q_2(t) S' \left(cx(t) - dy(t) + \rho^{(y)} \right) x \left(t + \frac{T}{2\pi} \chi \right) dt \quad . \end{aligned}$$

Here $S'(x) \equiv dS/du|_{u=x}$, Q_1 and Q_2 are the two components of \mathbf{Q} , and $x \left(t + \frac{T}{2\pi} \chi \right)$ and $y \left(t + \frac{T}{2\pi} \chi \right)$ are the two components of $\mathbf{\Gamma} \left(t + \frac{T}{2\pi} \chi \right)$, respectively. Following the argument of Kuramoto [31, 32], we keep only the sinusoidal terms of a Fourier series expansion of the right hand side of Eq. (S17), and obtain

$$\frac{d\phi_i}{dt} = \omega_i + \sum_{j \neq i} K_A A_{ij} \sin(\phi_j - \phi_i) + \sum_{j \neq i} K_A^{(\delta)} A_{ij}^{(\delta)} \sin(\phi_j - \phi_i - \delta) \quad , \quad (\text{S18})$$

where $A_{ij}^{(\delta)} \equiv B_{ij} + C_{ij}$, and K_A and $K_A^{(\delta)}$ are constants. Finally, we set

$$K_A \langle k \rangle = K_A^{(\delta)} \langle k^{(\delta)} \rangle \equiv K \quad ,$$

where $\langle k \rangle \equiv 1/N \sum_{i,j} A_{ij}$ and $\langle k^{(\delta)} \rangle \equiv 1/N \sum_{i,j} A_{ij}^{(\delta)}$ are the average degrees of the excitatory-excitatory and excitatory-inhibitory networks, respectively. This ensures that the two interaction terms on the right hand side of Eq. (S18) are of the same order. With this notation we obtain the two-layer MKM, i.e. Eq. (6),

$$\frac{d\phi_i}{dt} = \omega_i + \frac{K}{\langle k \rangle} \sum_{j \neq i}^N A_{ij} \sin(\phi_j - \phi_i) + \frac{K}{\langle k^{(\delta)} \rangle} \sum_{j \neq i}^N A_{ij}^{(\delta)} \sin(\phi_j - \phi_i - \delta) \quad . \quad (\text{S19})$$

THE CASE $\delta = 0$

In this section, we investigate some of the properties of the single-layer Kuramoto model,

$$d\phi_i/dt = \omega_i + K/\langle k \rangle \sum_{j=1}^N \alpha_{ij} \sin(\phi_j - \phi_i) \quad , \quad (\text{S20})$$

that is obtained by setting $\delta = 0$ in Eq. (6), and introducing the abbreviation $\alpha_{ij} \equiv A_{ij} + A_{ij}^{(\delta)} \langle k \rangle / \langle k^{(\delta)} \rangle$.

Synchronization. For sufficiently large N and sufficiently dense networks A_{ij} and $A_{ij}^{(\delta)}$, we expect the following behavior [33, 34]: For small values of K , the system attains a stationary global state that shows no synchronization, $\bar{r} \sim 0$, up to finite-size fluctuations of order $\mathcal{O}(1/\sqrt{N})$. Above a critical value,

$$K_c \equiv K_c^0 \langle k \rangle \frac{\langle \kappa \rangle}{\langle \kappa^2 \rangle} \quad , \quad (\text{S21})$$

synchronized clusters appear, $\bar{r} > 0$, where $K_c^0 \equiv 2/(\pi g(0))$, is the critical value for a fully connected network, and $\langle \kappa \rangle$ and $\langle \kappa^2 \rangle$ are the first two moments of the distribution of degrees $\kappa_i \equiv \sum_{j=1}^N \alpha_{ij}$, respectively. If both A_{ij} and $A_{ij}^{(\delta)}$ are Erdős-Rényi networks with probability p , κ_i follows a binomial distribution with probability $2p$. Hence we get $K_c = K_c^0/2 (1 - (1 - 2p)/2pN) + \mathcal{O}(N^{-2})$, meaning that the Kuramoto transition occurs at slightly lower values than for fully connected networks, as long as $\log(N)/N < p < 0.5$. For $K \gg K_c$, global synchronization is attained, $\bar{r} \sim 1$, after a transient time $\propto \langle k \rangle / (K \lambda_2)$, where λ_2 is the second-smallest eigenvalue of the Laplacian, $L_{ij} \equiv \kappa_i \delta_{ij} - \alpha_{ij}$.

Chaotic dynamics. The largest Lyapunov exponent takes the value $\bar{\lambda}_{max} = 0$ in the subcritical regime, and becomes positive around $K = K_c$, reaching peak values of $\bar{\lambda}_{max} \approx 0.05$. For $K > K_c$, $\bar{\lambda}_{max}$ decreases to marginally negative values $\bar{\lambda}_{max} \approx -0.001$, [35].

Average frequency deviation. For large values of K , solutions of Eq. (S20) are globally synchronized, and $\Omega = 1/N \sum_{i=1}^N \omega_i \approx \omega_0$. Related models [36–38] suggest that small phase shifts $\delta > 0$ do not destroy synchronization, so we expect $\Omega \approx \omega_0 - K/(\langle k^{(\delta)} \rangle N) \sum_{i,j} A_{ij}^{(\delta)} \cos(\phi_j - \phi_i) \sin(\delta) \approx \omega_0 - K\delta$, i.e. frequency suppression, for large values of K and small values of δ . Here we used properties of the sine function and of the synchronized solution.

CALCULATION OF THE INSTANTANEOUS LARGEST LYAPUNOV EXPONENT

In this section, we describe the algorithm that we use to calculate estimates for the instantaneous largest Lyapunov exponent $\lambda_{max}(t)$ together with the numerical integration of Eq. (6). Since solutions of Eq. (6) are vectors of phases $\phi = (\phi_1, \dots, \phi_N)$, an appropriate distance function $d(\phi, \psi)$ on the N -torus, measuring the separation between reference orbits and perturbed orbits at each time step, is needed. We choose the 1-norm distance

$$d(\phi, \psi) = \sum_{i=1}^N |\phi_i - \psi_i| \quad , \quad (\text{S22})$$

since it maximizes all p -norm distances, thus providing an upper bound on the separation distance. Note that making use of the embedding into Euclidean space $\phi \rightarrow (\sin(\phi), \cos(\phi))$ together with applying the 2-norm runs into obvious problems in the renormalization step of the algorithm (see below).

An initial separation between the reference trajectory and the perturbed trajectory of $d_0 = 10^{-4}$ is chosen. Both trajectories are integrated for one time step $dt = 0.1$ according to Eq. (6), and the quantity $\lambda_1 = \ln(d_1/d_0)/dt$ is calculated. d_1 is the separation between the trajectories after one step. For small initial separations d_0 this quantity is a good approximation for the instantaneous Lyapunov exponent after one time step [44]. Now the perturbed trajectory is renormalized in the direction of the difference vector to the reference trajectory to have a separation of d_0 again. After another integration step the quantity $\lambda_2 = \ln(d_2/d_0) + \ln(d_1/d_0)/(2dt)$ is calculated and the perturbed trajectory is renormalized as before. Iteration yields estimates for the instantaneous largest Lyapunov exponents

$$\lambda_{max}(t_n) = \lambda_n = \frac{1}{t_n - t_0} \sum_{i=1}^n \ln \left(\frac{d_i}{d_0} \right) \quad . \quad (\text{S23})$$

The algorithm was tested in a discrete version with the logistic map; $d_0 = 10^{-4}$ turned out to be sufficiently small to provide satisfactory results.